Scattering Theory of Nonlinear Thermoelectricity in Quantum Coherent Conductors

Jonathan Meair[1] and Philippe Jacquod[1,2]

[1]*Physics Department, University of Arizona, Tucson, AZ 85721, USA*
[2]*College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA*

We construct a scattering theory of weakly nonlinear thermoelectric transport through sub-micron scale conductors. The theory incorporates the leading nonlinear contributions in temperature and voltage biases to the charge and heat currents. Because of the finite capacitances of sub-micron scale conducting circuits, fundamental conservation laws such as gauge invariance and current conservation require special care to be preserved. We do this by extending the approach of Christen and Büttiker [Europhys. Lett. **35**, 523 (1996)] to coupled charge and heat transport. In this way we write relations connecting nonlinear transport coefficients in a manner similar to Mott's relation between the linear thermopower and the linear conductance. We derive sum rules that nonlinear transport coefficients must satisfy to preserve gauge invariance and current conservation. We illustrate our theory by calculating the efficiency of heat engines and the coefficient of performance of thermoelectric refrigerators based on quantum point contacts and resonant tunneling barriers. We identify in particular rectification effects that increase device performance.

I. INTRODUCTION

Today's energy challenges are calling for novel technologies to produce, store and use energy[1]. These challenges have among others renewed the interest in the old problem of thermoelectricity[2]. Research in thermoelectricity is predominantly directed towards finding new materials with better thermoelectric properties[3–5], however a recently proposed alternative approach has been to try and tailor low-dimensional micro- and nanostructured materials[6], which may exhibit the sharp oscillations in the electronic density of states necessary to trigger large thermoelectric effects[7]. With the ongoing trend towards circuit miniaturization in electronics, heat evacuation becomes a major problem and it is accordingly important to explore how heat currents can be controlled at sub-micron scales. Thus, there is ample motivation for investigating thermoelectricity in micro- and nanodevices[8]. To name but few examples of recent investigations, sub-micron electrically controlled heat rectifiers[9], thermally driven electric rectifiers[10,11], thermoelectric refrigerators[12–14] and heat converters[15–17] have been proposed, and thermoelectric effects in molecular transport have been investigated[18–24]. On a more fundamental level, thermoelectric effects provide information on the nature of the underlying quantum fluid, for instance, breakdowns of the Wiedemann-Franz law[25] may reflect non Fermi liquid behaviors[26–28], quantum interference effects[29,30] or electric and heat transport mediated by both electron and hole quasiparticles in Andreev systems[31–33].

The field of thermoelectricity has until now overlooked the fact that thermoelectric devices often operate in the nonlinear regime of transport (See for instance Refs. [18,34,35]). Rectification effects occur in that regime[36–40], and their influence on thermoelectricity is not yet understood. There are only few works we know about which have discussed nonlinear thermoelec-

tricity. Ref. [41] calculated the thermopower to quadratic order in temperature gradients and Refs. [42–44] focused on the nonlinear Peltier coefficient. These works considered bulk systems with large capacitances, where accordingly the build-up of local electrostatic potentials can be neglected. This neglect is no longer justified in small constrictions, where such potentials have to be self-consistently connected to the applied biases in order to preserve gauge invariance[45]. Refs. [46,47] discussed nonlinear thermoelectricity in such constrictions but did not introduce local potentials, thereby taking the risk of violating gauge invariance. Refs. [14,34,48] introduced self-consistently determined homogeneous Hartree potentials to preserve gauge invariance.

In this manuscript we generalize the scattering theory of weakly nonlinear transport of Ref. [45] to coupled electric and heat transport. We go beyond Ref. [48] which considered only electric currents in the presence of voltage and temperature biases, by additionally treating heat currents. One of our motivations is to provide criteria for enhanced thermoelectric efficiency in the nonlinear regime. The standardly used figure of merit ZT is constructed out of linear thermoelectric coefficients only[2], thus while any measure of thermodynamic efficiency in the linear regime can only be a function of ZT, it is expectable that ZT loses its predictive power in the nonlinear regime. It is often the case, and we illustrate this below, that the maximal efficiency predicted by ZT is attained outside the linear regime. In this article we accordingly focus our attention on thermodynamic efficiencies and coefficients of performance as indicators of thermoelectric performance. In particular we investigate how they behave upon increasing sources of nonlinearities such as temperature and voltage biases and capacitive couplings. A key aspect of our theory is that in the nonlinear regime of transport, the state of the system is self-consistently determined by the conditions under which transport proceeds. We will take this self-consistency

into account by introducing piecewise constant local electrostatic potentials. We will go beyond Refs. [14,34,48] and consider more than one such potential. In the scattering theory, gauge invariance and current conservation are naturally preserved in the linear regime by the unitarity of the scattering matrix. In the nonlinear regime, they are reflected by sum rules connecting nonlinear transport coefficients. Such sum rules were derived in Ref. [45] for purely electric transport and below we give a complete list of sum rules ensuring gauge invariance and current conservation for weakly nonlinear thermoelectric transport.

The paper is organized as follows. In Section II, we construct a scattering theory of weakly nonlinear thermoelectric transport valid to quadratic order in temperature and voltage biases. We express linear and nonlinear transport coefficients in terms of the scattering matrix and the local potentials generated in the system by the biases. In Section III, we discuss these local potentials and recall how they can be determined in practice. In Section IV we write down sum rules that have to be satisfied for the theory to preserve charge and current conservation as well as gauge invariance. In Section V we briefly discuss performance metrics for thermoelectric devices. In Section VI we apply our theory to two particular devices, a quantum point contact and a resonant tunneling double-barrier. We find in particular that the bias range of operation of these devices is such that our weakly nonlinear approximation gives a rather accurate description of the exact thermodynamic efficiency. Conclusions are presented in Section VII.

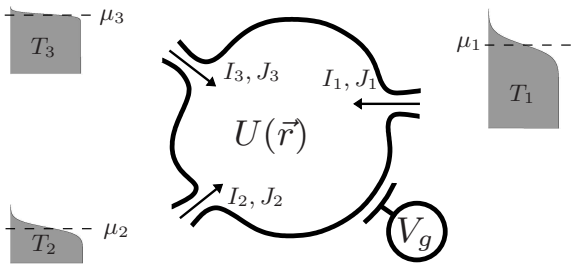## II. SCATTERING APPROACH TO THERMOELECTRICITY

FIG. 1: Sketch of a multi-terminal coherent conductor capacitively coupled to an external gate and connected to three external Fermi liquid reservoirs. The electronic Fermi distributions of the reservoirs are depicted. In the nonlinear regime, an inhomogeneous electrostatic potential $U(\vec{r})$ develops inside the conductor, as it is induced by the applied voltages and temperatures at the reservoirs. The arrows indicate our convention that electric and heat currents are positive when they flow into the conductor.

In linear response, the scattering theory to thermoelectric transport was worked out by Butcher[50]. We follow

his approach and consider heat and electric transport mediated only by electron quasiparticles and extend it to the nonlinear regime. Considering a coherent conductor connected to $i = 1, 2, ...M$ external reservoirs via reflectionless leads (see Fig. 1), the electric (I_i) and heat (J_i) currents through contact i are given by

$$I_i = \frac{2e}{h} \sum_j \int d\varepsilon \ f_j(\varepsilon) \ A_{ij} \left[\varepsilon, U(\vec{r})\right], \tag{1a}$$

$$J_i = \frac{2}{h} \sum_j \int d\varepsilon \ (\varepsilon - eV_i) \ f_j(\varepsilon) \ A_{ij} \left[\varepsilon, U(\vec{r})\right]. \tag{1b}$$

We took the convention that both electric and heat currents are positive when they enter the conductor from a reservoir. Thus, for example, the conductor is cooled if $\sum_i J_i < 0$ and heated otherwise. In Eqs. (1), $\varepsilon \equiv E - E_F$ is the energy of an electron counted from the Fermi energy, $f_j(\varepsilon) = \{1 + \exp\left[(\varepsilon - eV_j)/k_B T_j\right]\}^{-1}$ is the Fermi function in reservoir j, maintained at a voltage V_j and a temperature T_j, $A_{ij}\left[\varepsilon, U(\vec{r})\right] = N_i[\varepsilon, U(\vec{r})]\delta_{i,j} - \mathcal{T}_{ij}\left[\varepsilon, U(\vec{r})\right]$ is given by the number N_i of transport channel in lead i and the scattering matrix $\mathbf{s}$ via $\mathcal{T}_{ij} = \mathrm{Tr}[\mathbf{s}_{ij}^\dagger \mathbf{s}_{ij}]$ where $\mathbf{s}_{ij}$ is the sub-block of $\mathbf{s}$ connecting terminals j to i and $U(\vec{r})$ is the local potential in the conductor. In this manuscript, we take N_i to be independent of energy and biases. The electrostatic potential induced within the conductor (see Fig. 1) is itself dependent on the voltages and temperatures in the reservoirs, i.e., $U(\vec{r}, \{V_k\}, \{T_k\})$. Expanding the currents in Eqs. (1) to quadratic order in temperature differences, $\theta_j \equiv T_j - T_0$, and voltage biases one obtains

$$
\begin{aligned}
I_i =& \sum_j (G_{ij} V_j + B_{ij} \theta_j) \\
&+ \sum_{j,k} [G_{ijk} V_j V_k + B_{ijk} \theta_j \theta_k + Y_{ijk} V_j \theta_k], \tag{2a}
\end{aligned}
$$

$$
\begin{aligned}
J_i =& \sum_j (\Gamma_{ij} V_j + \Xi_{ij} \theta_j) \\
&+ \sum_{j,k} [\Gamma_{ijk} V_j V_k + \Xi_{ijk} \theta_j \theta_k + \Psi_{ijk} V_j \theta_k]. \tag{2b}
\end{aligned}
$$

The linear response coefficients were calculated in Ref. [50]. They are given by

$$G_{ij} = \frac{2e^2}{h} \int d\varepsilon \ (-\partial_\varepsilon f) \ A_{ij}(\varepsilon), \tag{3a}$$

$$B_{ij} = \frac{2e}{hT_0} \int d\varepsilon \ (-\partial_\varepsilon f) \ \varepsilon \ A_{ij}(\varepsilon), \tag{3b}$$

$$\Gamma_{ij} = \frac{2e}{h} \int d\varepsilon \ (-\partial_\varepsilon f) \ \varepsilon \ A_{ij}(\varepsilon), \tag{3c}$$

$$\Xi_{ij} = \frac{2}{hT_0} \int d\varepsilon \ (-\partial_\varepsilon f) \ \varepsilon^2 \ A_{ij}(\varepsilon), \tag{3d}$$

where $f(\varepsilon) = \{1 + \exp\left[\varepsilon/k_B T_0\right]\}^{-1}$. The temperature $T_0 \neq 0$ is in principle arbitrary, however, better convergence of the expansion of Eqs. (2) is obtained when T_0

is taken as the temperature in one of the reservoirs or the average of the reservoir temperatures. In particular, the sum rules given below in Section IV guarantee that all currents I_i and J_i vanish at equilibrium, regardless of the choice of T_0. Eqs. (3b) and (3c) show that the reciprocity relation $\Gamma_{ij} = B_{ij}T_0$ is automatically satisfied (see Ref. [51] for a detailed discussion of Onsager relations in the linear regime).

These linear response coefficients are determined by the equilibrium state of the system for which we set $U^{(eq)}(\vec{r}) = 0$. However, in the nonlinear regime, we must consider how the biases affect the local potential. We do that by introducing characteristic potentials[45,48]

$$U(\vec{r}) = \sum_k \left[u_k(\vec{r})\, V_k + z_k(\vec{r})\, \theta_k \right], \quad (4)$$

which describe how $U(\vec{r})$ depends on the biases to leading order. Truncating the expansion of $U(\vec{r})$ to linear order in the biases is sufficient to correctly capture currents to quadratic order. Gauge invariance requires that $\sum_k u_k(\vec{r}) = 1$[45], however there is no such relation involving $z_k(\vec{r})$. The nonlinear response coefficients are then

$$G_{ijk} = \frac{e^3}{h} \int d\varepsilon \, (-\partial_\varepsilon f)\, \alpha_{ijk}(\varepsilon)\,, \quad (5a)$$

$$B_{ijk} = \frac{e}{hT_0^2} \int d\varepsilon \, (-\partial_\varepsilon f)\, \varepsilon\, \beta_{ijk}(\varepsilon)\,, \quad (5b)$$

$$Y_{ijk} = \frac{2e^2}{hT_0} \int d\varepsilon \, (-\partial_\varepsilon f)\, \gamma_{ijk}(\varepsilon)\,, \quad (5c)$$

$$\Gamma_{ijk} = \frac{e^2}{h} \int d\varepsilon \, (-\partial_\varepsilon f)\, \varepsilon\, \alpha_{ijk}(\varepsilon)$$
$$- G_{ik}\left(\delta_{i,j} - \frac{1}{2}\delta_{j,k} \right)\,, \quad (5d)$$

$$\Xi_{ijk} = \frac{1}{hT_0^2} \int d\varepsilon \, (-\partial_\varepsilon f)\, \varepsilon^2\, \beta_{ijk}(\varepsilon) + \frac{\Xi_{ij}}{2T_0}\delta_{j,k}\,, \quad (5e)$$

$$\Psi_{ijk} = \frac{2e}{hT_0} \int d\varepsilon \, (-\partial_\varepsilon f)\, \varepsilon\, \gamma_{ijk}(\varepsilon)$$
$$- B_{ik}\left(\delta_{i,j} - \delta_{j,k} \right)\,, \quad (5f)$$

where we defined

$$\alpha_{ijk}(\varepsilon) = \int d\mathbf{r} \frac{\delta A_{ij}(\varepsilon)}{e\delta U(\vec{r})} \left[2u_k(\vec{r}) - \delta_{j,k} \right]\,, \quad (6a)$$

$$\beta_{ijk}(\varepsilon) = \int d\mathbf{r} \frac{\delta A_{ij}(\varepsilon)}{e\delta U(\vec{r})} \left[2eT_0 z_k(\vec{r}) - \varepsilon\delta_{j,k} \right]\,, \quad (6b)$$

$$\gamma_{ijk}(\varepsilon) = \int d\mathbf{r} \left(\frac{\delta A_{ij}(\varepsilon)}{e\delta U(\vec{r})} \left[eT_0 z_k(\vec{r}) - \varepsilon\delta_{j,k} \right] \right.$$
$$\left. + \frac{\delta A_{ik}(\varepsilon)}{e\delta U(\vec{r})}\, \varepsilon\, u_j(\vec{r}) \right)\,. \quad (6c)$$

Several of the nonlinear response coefficients of Eqs. (5) differ primarily by a factor of ε inside the energy integral. Using a Sommerfeld expansion, we find that at low temperature the following relations between nonlinear response coefficients must hold,

$$\frac{\Gamma_{ijk} + G_{ik}(\delta_{i,j} - \delta_{j,k}/2)}{G_{ijk}} = \frac{(\pi k_B T_0)^2}{3e}\, \partial_\varepsilon \ln G_{ijk}\,, \quad (7a)$$

$$\frac{\Psi_{ijk} + B_{ik}(\delta_{i,j} - \delta_{j,k})}{Y_{ijk}} = \frac{(\pi k_B T_0)^2}{3e}\, \partial_\varepsilon \ln Y_{ijk}\,, \quad (7b)$$

$$\frac{B_{ijk}}{\Xi_{ijk} - \Xi_{ij}\delta_{j,k}/(2T_0)} = e\, \partial_\varepsilon \ln \left[\Xi_{ijk} - \frac{\Xi_{ij}\delta_{j,k}}{2T_0} \right]\,, \quad (7c)$$

where all derivatives on the right-hand side are evaluated at the Fermi energy, $\varepsilon = 0$. These relations are the nonlinear counterpart to the Mott relation[25] $S = -(\pi^2 k_B^2 T_0/3e)\partial_\varepsilon \ln G$ between the thermopower $S = -B_{12}/G_{12}$ and the logarithmic derivative of the conductance $G \equiv G_{12}$ in two-terminal conductors[52]. There is no counterpart to the Onsager reciprocity relations for nonlinear coefficients, because, as can be seen from Eqs. (5), they all have different energy integrands. The absence of such relations may be attributed to the breaking of microscopic reversibility in the nonlinear regime, where local potentials are self-consistently related to transport, and where their dependence on temperature [the coefficients z_k in Eq. (4)] is not necessarily related to their dependence on voltages [the coefficients u_k in Eq. (4)].

III. CHARACTERISTIC POTENTIALS

Our main task is next to determine the characteristic potentials defined in Eq. (4). We briefly summarize how this can be done[45,48]. We consider first how additional charges are injected into the conductor as it is driven out of equilibrium. Biased reservoirs inject a net addition of charge, called the bare charge, $\delta q^{(b)}$. The presence of $\delta q^{(b)}$ induces an electrostatic potential shift within the conductor which, in its turn, generates a screening charge $\delta q^{(s)}$. The net injected charge is the sum of the two, $\delta q(\vec{r}) = \delta q^{(b)}(\vec{r}) + \delta q^{(s)}(\vec{r})$. To leading order the bare charge can be expressed as a linear function of the biases,

$$\delta q^{(b)}(\vec{r}) = \sum_j D_j^V(\vec{r})V_j + \sum_j D_j^\theta(\vec{r})\theta_j, \quad (8)$$

where $D_j^V(\vec{r}) = e^2 \int d\varepsilon(-\partial_\varepsilon f)\ \nu_j^p(\vec{r},\varepsilon)$ and $D_j^\theta(\vec{r}) = e \int d\varepsilon(-\partial_\varepsilon f)\ \nu_j^e(\vec{r},\varepsilon)$ are *charge injectivities* (expressed in the units of capacitances as in Ref. [45]) determined by the energy dependent *particle*[45] and *entropic*[48] *injectivities*

$$\nu_j^p(\vec{r},\varepsilon) = -\frac{1}{4\pi i} \sum_k \text{Tr}\left[\mathbf{s}_{kj}^\dagger \frac{\delta \mathbf{s}_{kj}}{e\delta U(\vec{r})} - \frac{\delta \mathbf{s}_{kj}^\dagger}{e\delta U(\vec{r})} \mathbf{s}_{kj} \right]\,, \quad (9a)$$

$$\nu_j^e(\vec{r},\varepsilon) = \left(\frac{\varepsilon}{T_0} \right) \nu_j^p(\vec{r},\varepsilon)\,, \quad (9b)$$

where the trace runs over channel indices. Eq. (9a) gives the partial density of states at $\vec{r}$ corresponding to injection from lead j, which is multiplied by a single-particle entropy factor ε/T_0 in Eq. (9b)[48].

Next, the screening charge $\delta q^{(s)}(\vec{r})$ is related to $U(\vec{r})$ via the Lindhard polarization function. The relation is in principle nonlocal, however the nonlocal corrections are quantum in essence and thus negligible in good metals with high enough electronic density. We thus make the quasiclassical approximation of a local Lindhard function and write $\delta q^{(s)}(\vec{r}) = -\Pi(\vec{r})U(\vec{r})$, with $\Pi(\vec{r}) = \sum_j D_j^V(\vec{r}) = D(\vec{r})$. From the definition of D_j^V we have $D(\vec{r}) = e^2 \int d\varepsilon(-\partial_\varepsilon f)\, \nu(\vec{r}, \varepsilon)$ with the local density of states $\nu(\vec{r}, \varepsilon) = \sum_j \nu_j^p(\vec{r}, \varepsilon)$. Note that at zero temperature, $D(\vec{r})$ is the density of states (in units of capacitance) at the Fermi energy. Putting all this together, the net total injected charge is, to leading order in the biases,

$$\delta q(\vec{r}) = \sum_j D_j^V(\vec{r})V_j + \sum_j D_j^\theta(\vec{r})\theta_j - D(\vec{r})U(\vec{r}). \quad (10)$$

The conductor may also be capacitively coupled to the reservoirs and external metallic gates. One way to take these couplings into account is to discretize the conductor into different regions labelled $r = 1, 2, ...,$ each with a constant potential U_r. The procedure brings about additional relations

$$\delta q_r = \sum_k C_{rk}(U_r - V_k), \quad (11)$$

with the sum running over all capacitively coupled components. This equation expresses changes in local charge distributions due to Coulomb interactions. The latter are represented here by geometric capacitances between different components of the system.

Once the microstructure considered is consistently defined, in that all electric components (leads, gates, etc.) that are coupled to the conductor are included so that electric field lines do not leave the system, then the total charge is conserved by Gauss' theorem. Eqs. (4), (10) and (11) then determine the characteristic potentials in Eqs. (6). The nonlinear transport coefficients are finally calculated by using the identity $\int d\mathbf{r}\, \delta_{U(\vec{r})}(...) = -e\partial_E(...)$ to compute the functional derivatives in Eqs. (6). Two examples of thermoelectric microdevices will be discussed in Section VI, which illustrate how the procedure is applied in practice by discretizing the microstructures into regions of uniform local potential.

IV. CONSERVATIONS AND SUM RULES

Particle conservation is expressed mathematically by the unitarity of the scattering matrix, $\sum_i A_{ij} = \sum_j A_{ij} = 0$. This translates into the following sum rules for the linear transport coefficients,

$$\sum_{i\ \text{or}\ j} G_{ij} = \sum_{i\ \text{or}\ j} B_{ij} = \sum_{i\ \text{or}\ j} \Gamma_{ij} = \sum_{i\ \text{or}\ j} \Xi_{ij} = 0. \quad (12)$$

In particular, two of these sum rules correspond to electric current conservation, $\sum_i G_{ij} = \sum_i B_{ij} = 0$ while two others reflect gauge invariance, $\sum_j G_{ij} = \sum_j \Gamma_{ij} = 0$.

The unitarity of the scattering matrix further implies $\sum_i \alpha_{ijk} = \sum_i \beta_{ijk} = \sum_i \gamma_{ijk} = \sum_j (\alpha_{ijk} + \alpha_{ikj}) = \sum_j \gamma_{ijk} = 0$, which results in the following sum rules between the nonlinear response coefficients,

$$\sum_i G_{ijk} = \sum_i B_{ijk} = \sum_i Y_{ijk} = 0, \quad (13a)$$

$$\sum_j (G_{ijk} + G_{ikj}) = \sum_j (\Gamma_{ijk} + \Gamma_{ikj}) = 0, \quad (13b)$$

$$\sum_j Y_{ijk} = \sum_j \Psi_{ijk} = 0, \quad (13c)$$

$$\sum_i \Xi_{ijk} = 0, \quad (13d)$$

$$G_{jk} = -\sum_i \Gamma_{ijk}, \qquad B_{jk} = -\sum_i \Psi_{ijk}. \quad (13e)$$

Two of these sum rules were derived in Ref. [45], $\sum_i G_{ijk} = 0$, implied by current conservation and $\sum_j (G_{ijk} + G_{ikj}) = 0$, resulting from gauge invariance. All other sum rules are written here for the first time. Eqs. (13a) reflect current conservation while Eqs. (13b) and (13c) correspond to gauge invariance. The sum rules of Eq. (13e) imply $\sum_i J_i + I_i^{(\text{lin})}V_i = 0$, where $I_i^{(\text{lin})}$ is the linear response current through lead i. Thus in our approach energy is conserved to quadratic order in the biases.

Finally we note the following, additional double-sum rules $\sum_{jk} \alpha_{ijk} = \sum_{jk} \beta_{ijk} = 0$ which imply

$$\sum_{j,k} G_{ijk} = \sum_{j,k} \Gamma_{ijk} = 0, \quad (14a)$$

$$\sum_{j,k} B_{ijk} = \sum_{j,k} \Xi_{ijk} = 0. \quad (14b)$$

While Eqs. (14a) are guaranteed to hold by Eqs. (13b), Eqs. (14b) are new. Eqs. (12), (13c), (14a) and (14b) ensure in particular that the electric and heat currents of Eqs. (2) vanish at equilibrium.

V. PERFORMANCE METRICS FOR THERMOELECTRICITY

For two-terminal setups treated in linear response, Eqs. (1) reduce to

$$\begin{pmatrix} I \\ J \end{pmatrix} = \begin{pmatrix} G & B \\ \Gamma & \Xi \end{pmatrix} \begin{pmatrix} V \\ \Delta T \end{pmatrix}, \quad (15)$$

where V and ΔT are the voltage and temperature differences between the reservoirs. Alternatively, this linear relation is often rewritten in terms of responses to an electric current and a temperature difference,

$$\begin{pmatrix} V \\ J \end{pmatrix} = \begin{pmatrix} R & S \\ \pi_p & \kappa \end{pmatrix} \begin{pmatrix} I \\ \Delta T \end{pmatrix}. \quad (16)$$

Eqs. (15) and (16) define the linear response coefficients, the electric conductance, G, and resistance, R, the Seebeck, S, and Peltier, π_p, coefficients as well as the thermal conductances Ξ (at zero voltage, $V = 0$) and κ (at zero current, $I = 0$). We can express the relationships between these linear transport coefficients as

$$
\begin{pmatrix} R & S \\ \pi_p & \kappa \end{pmatrix} = \begin{pmatrix} 1/G & -B/G \\ \Gamma/G & \Xi - B\Gamma/G \end{pmatrix}. \quad (17)
$$

Dimensional analysis shows that a single dimensionless quantity can be constructed from the linear response coefficients and the temperature. This is the figure of merit $ZT = (GS^2/\kappa)T_0$. In the linear regime, thermodynamic efficiencies are then only functions of ZT and the temperature difference, for instance a heat engine working in the linear regime has maximal efficiency[49]

$$
\eta_{\max} = \max_V \left(\frac{IV}{J} \right) = \left(\frac{\Delta T}{T_0} \right) \frac{\sqrt{ZT+1}-1}{\sqrt{ZT+1}+1}, \quad (18)
$$

It is important to keep in mind that this relation is valid only in the linear regime, and that nonlinear effects may invalidate it. Linear response may even predict a maximal efficiency to occur outside its range of validity (at too large a bias voltage). A description of thermodynamic efficiency based only on ZT is therefore inappropriate in the nonlinear regime. However, the efficiency is generically given by $\eta = IV/J_{\text{hot}}$, where J_{hot} is the heat current flowing from the hot reservoir. While $J_{\text{hot}} = -J_{\text{cold}}$ in linear response, $J_{\text{hot}} \neq -J_{\text{cold}}$ in the nonlinear regime. Note that with this definition, it makes sense to substitute $T_0 \to T_{\text{hot}}$ in Eq. (18), in which case one obtains Carnot efficiency for $ZT \to \infty$.

The efficiency of thermoelectric refrigeration is measured by the coefficient of performance (COP) which is the ratio of the heat extracted from the cold reservoir to the work done[2], $\text{COP} = J_{\text{cold}}/IV$. This refrigeration "efficiency" is limited by $\text{COP} < T_{\text{cold}}/\Delta T$. In the linear regime, its maximum is also given by ZT as

$$
\text{COP}_{\max} = \max_V \left(\frac{J}{IV} \right) = \left(\frac{T_0}{\Delta T} \right) \frac{\sqrt{ZT+1}-1}{\sqrt{ZT+1}+1}, \quad (19)
$$

where this time it makes sense to substitute $T_0 \to T_{\text{cold}}$. Note that the linear response scattering theory leaves the choice of T_0 arbitrary, up to the fact that a well chosen T_0 (close to both T_{cold} and T_{hot}) minimizes nonlinear corrections.

It is not at all guaranteed that ZT remains a valid figure of merit once nonlinear effects kick in. Below we discuss two examples of mesoscopic thermoelectric engines and compare their true efficiency to that predicted by linear response.

VI. EXAMPLES

A. Resonant Tunneling Barrier

We illustrate our theory, first by calculating the thermodynamic efficiency of thermoelectric devices based on a coherent single-level double-barrier system - a resonant tunneling barrier (RTB). Nonlinear electric transport through a RTB was investigated in Ref. [45], where rectification effects were stressed in asymmetric systems with different tunneling probabilities through the left and right barriers. Electric transport generated by simultaneously applied voltage and temperature biases was investigated in Ref. [48], focusing on the thermopower. This system is especially interesting in that the weakly nonlinear theory can be compared to fully nonlinear solutions[45].
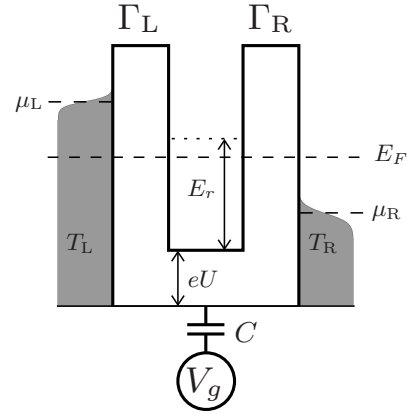

FIG. 2: Sketch of a resonant tunneling barrier (RTB) formed between two Fermi liquid reservoirs by two tunnel barriers in series with transparencies Γ_L and Γ_R. They define a potential well with a single resonant level capacitively coupled to an external gate.

The system we consider is sketched in Fig. 2. Two tunnel barriers with tunnel rates $\Gamma_{L,R}$ define a potential well with a single resonant level, at energy E_r above the bottom of the well. Left and right, the well is coupled to two Fermi liquid reservoirs at temperatures $T_{L,R}$ and electrochemical potentials $\mu_{L,R}$. The well is capacitively coupled to an external gate with voltage V_g which, together with the temperatures and electrochemical potentials in the reservoirs determine the potential U (assumed spatially homogeneous) in the well. This models a quantum dot in the Coulomb blockade regime as long as neither the temperatures nor the voltage biases exceed the single-particle level spacing in the dot. The equilibrium energy dependent particle injectivity at lead $j = $L, R is[45]

$$
\nu_j^p(\varepsilon) = \frac{1}{\pi} \frac{\Gamma_j}{(\varepsilon + E_F - E_r)^2 + \Gamma^2}, \quad (20)
$$

and the transmission probability is[45]

$$
\mathcal{T}(\varepsilon) = \frac{4\Gamma_L \Gamma_R}{(\varepsilon + E_F - E_r)^2 + \Gamma^2}, \quad (21)
$$

where $\Gamma = \Gamma_\mathrm{L} + \Gamma_\mathrm{R}$ is the total decay rate of the resonant level and we defined $E_F = (\mu_\mathrm{L} + \mu_\mathrm{R})/2$. In the following we will set $E_F = 0$, as the zero of the energy scale and measure all other energies in units of $\Gamma = \mathrm{const}$.

Because we consider a single homogeneous potential in the well, $U(\vec{r}) = U$, the injected charge loses its spatial dependence, $\delta q(\vec{r}) \to \delta q$. Consequently, Eq. (10) can be rewritten as

$$\delta q = \sum_j D_j^V V_j + \sum_j D_j^\theta \theta_j - DU . \quad (22)$$

The well is capacitively coupled to an external gate, thus Eq. (11) tells us that the additional charge in the potential well is also given by

$$\delta q = C(U - V_g) . \quad (23)$$

Using Eqs. (22) and (23) we obtain the characteristic potentials as

$$u_j = \frac{D_j^V}{C + D}, \quad u_g = \frac{C}{C + D}, \quad z_j = \frac{D_j^\theta}{C + D} , \quad (24)$$

where the charge and entropic injectivities are given by $D_j^V = e^2 \int d\varepsilon(-\partial_\varepsilon f) \; \nu_j^p(\varepsilon)$ and $D_j^\theta = e \int d\varepsilon(-\partial_\varepsilon f) \; (\varepsilon/T_0) \nu_j^p(\varepsilon)$, with the particle injectivity ν_j^p of Eq. (20). This fully determines the nonlinear coefficients in Eqs. (5) so that currents can be calculated using Eqs. (2).

For the RTB the potential U can alternatively be determined nonperturbatively. Eq. (22) for the injected
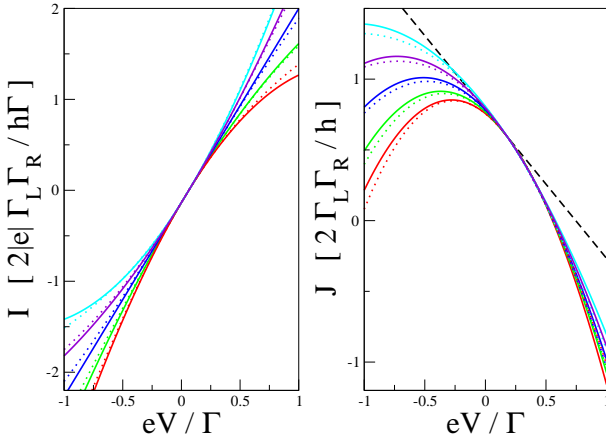

FIG. 3: Electric (left) and heat (right) currents from a hot reservoir to a RTB with $(E_r - E_F)/\Gamma = 1$, $\mu_\mathrm{L} - \mu_\mathrm{R} = eV$ and $T_\mathrm{R} - T_\mathrm{L} = 0.3 \, T_0$. The base temperatures are $k_B T_0/\Gamma = 0.3$ (left) and $k_B T_0/\Gamma = 1$ (right). The different curves correspond to different tunnel barrier asymmetries $(\Gamma_\mathrm{L} - \Gamma_\mathrm{R})/\Gamma = 0.95$ (red), 0.5 (green), 0 (blue), -0.5 (violet), -0.95 (cyan). Here, we switch off the capacitive coupling between the RTB and the external gate, $C = 0$. The solid lines are the approximate solutions in our weakly nonlinear approach and the dotted lines are the fully nonlinear solutions. The black dashed line is the heat current predicted from linear response.
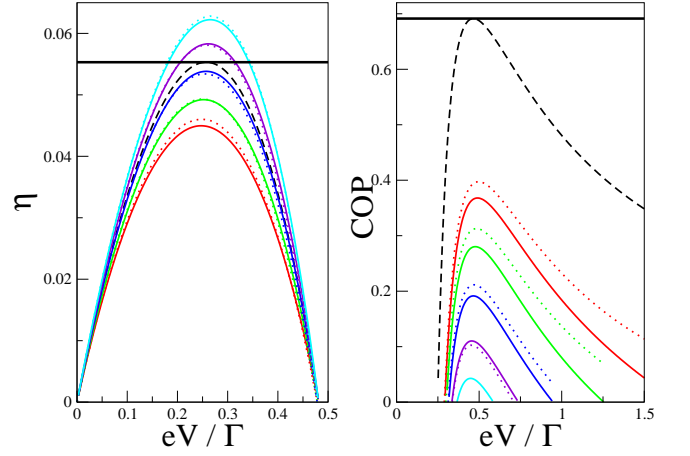

FIG. 4: Left panel: Thermodynamic efficiency for a RTB-based heat engine with $(E_r - E_F)/\Gamma = 1$, $\mu_\mathrm{L} - \mu_\mathrm{R} = eV$, $k_B T_0/\Gamma = 1$, $T_\mathrm{R} - T_\mathrm{L} = 0.8 \, T_0$ and different asymmetries $(\Gamma_\mathrm{L} - \Gamma_\mathrm{R})/\Gamma = 0.95$ (red), 0.5 (green), 0 (blue), -0.5 (violet), -0.95 (cyan). Right panel: COP for the same RTB as in the left panel with $T_\mathrm{R} - T_\mathrm{L} = 0.1 \, T_0$ and working at larger bias, thus as a refrigerator. In both cases there is no capacitive coupling between the RTB and the external gate, $C = 0$. The solid lines are the weakly nonlinear approximate solutions and the dotted lines are the fully nonlinear solutions (see text). The dashed black line is the linear response solution and the solid black line is the maximum efficiency/COP predicted from linear response.

charge is replaced by the energy integral

$$\delta q = e \int d\varepsilon \left[\sum_j \nu_j^p(\varepsilon - eU) \, f_j(\varepsilon) - \nu(\varepsilon) \, f(\varepsilon) \right] , \quad (25)$$

with $\nu(\varepsilon) = \sum_j \nu_j^p(\varepsilon)$. One determines U numerically from Eqs. (23) and (25). The fully nonlinear currents are then obtained from Eqs. (1). The result is exact, up to the approximation of a single homogeneous local potential in the well.

We first show in Fig. 3 the electric and heat currents for two RTB's. Different temperatures have been chosen to present I and J in order to visually optimize nonlinear effects. The curves are clearly nonlinear and display rectification in both the electric (left) and heat (right) currents. The rectification in the heat current is in general more pronounced, due to the additional voltage in the integrand of Eq. (1b). Most remarkably, perhaps, we see that the behavior obtained via the weakly nonlinear approximation (solid lines) does not differ much from the fully nonlinear result (dotted lines), even for the strongly nonlinear heat current. The pronounced rectification effects will turn out to boost or hinder the efficiency of RTB-based devices depending on the asymmetry $(\Gamma_\mathrm{L} - \Gamma_\mathrm{R})/\Gamma$.

We next investigate the efficiency of RTB-based thermoelectric devices. Fig. 4 shows the thermodynamic efficiency, $\eta = IV/J_\mathrm{hot}$, of a heat engine (left panel) and

the coefficient of performance, COP $= J_{\mathrm{cold}}/IV$, of a refrigerator in the linear (dashed black line) and in the nonlinear approximation (colored solid and dotted lines) for different asymmetries $(\Gamma_{\mathrm{L}} - \Gamma_{\mathrm{R}})/\Gamma$. In linear response, $I \propto \Gamma_{\mathrm{L}}\Gamma_{\mathrm{R}}/\Gamma$ and $J \propto \Gamma_{\mathrm{L}}\Gamma_{\mathrm{R}}$, so that both η and COP are independent of $(\Gamma_{\mathrm{L}} - \Gamma_{\mathrm{R}})/\Gamma$. For the chosen parameters, the maximal efficiency is at finite voltage bias, yet linear response gives a good approximation to η in the symmetric case, $\Gamma_{\mathrm{L}} = \Gamma_{\mathrm{R}}$. This is so, because in that case, $D_{\mathrm{L}}^{V,\theta} = D_{\mathrm{R}}^{V,\theta}$ and the screening potential given by Eq. (22) identically vanishes in the absence of gate coupling, $C = 0$. Introducing an asymmetry in the RTB changes the rules of the game, as rectification in the electric and heat currents kick in. As the asymmetry increases we can see a clear deviation from linear response, which emphasizes the importance of treating screening effects when characterizing thermoelectric performances of such microdevices. The efficiency η increases or decreases, depending on whether rectification play with or against the bias voltage. This depends on the sign of the bias voltage and the direction of rectification. From Fig. 3 we see that at positive voltage, samples with large negative values for the asymmetry $(\Gamma_{\mathrm{L}} - \Gamma_{\mathrm{R}})/\Gamma$ exhibit rectification effects where simultaneously the electric current is enhanced while the heat current is reduced compared to their values in the linear regime. Both effects conspire to increase the efficiency in that case (cyan curves in the left panel of Fig. 4). The coefficient of performance, on the other hand, is always significantly smaller than predicted by linear response, and rectification plays in the opposite direction in the refrigerator case because COP $= J_{\mathrm{cold}}/IV$ while $\eta = IV/J_{\mathrm{hot}}$. Another remarkable feature is that both performance metrics are well predicted by the weakly nonlinear theory (compare solid with dotted lines) with larger deviations, as expected, in the refrigerator case at larger biases.

So far the capacitive coupling to the external gate has been set equal to zero, $C = 0$. Fig. 5 finally shows the dependence of η and COP on C. Efficiencies and maximal efficiencies for a heat engine are shown in the left panels, while coefficients of performance and maximal coefficients of performance for a refrigerator are on the right panels. For the chosen set of parameters, linear response overestimates η and COP when there is no coupling to the gate. In both panels c) and d), the capacitive coupling is seen to increase η and COP. Quite remarkably the efficiency ranges from well below the linear prediction for $C\Gamma/e^2 = 0$ to well above it for $C\Gamma/e^2 = 10^{-5}$, strikingly illustrating the dependence of η on nonlinear effects. The same calculation for RTB's with all other asymmetries reported in Fig. 4 show systematic increases in η above its value at $C\Gamma/e^2 = 0$ as the capacitive coupling is turned on. The magnitude of the increase is in all cases similar to that in Fig. 5, which we attribute to similar increases in U with C, regardless of the asymmetry. Panels a) and b) also show that better performances are obtained for a resonant level lying a distance of a few times Γ above or below E_F. We finally note that, as

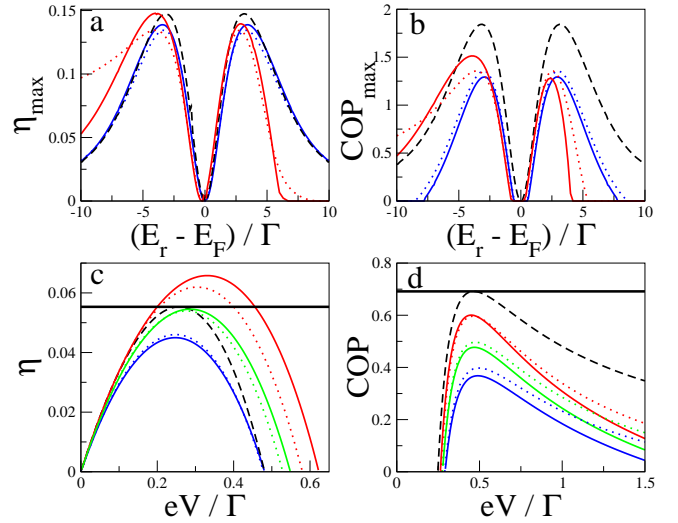FIG. 5: a) Maximal thermodynamic efficiency, $\eta_{\mathrm{max}} = \max_V \eta$, for a RTB heat engine and b) maximal coefficient of performance, $\mathrm{COP}_{\mathrm{max}} = \max_V \mathrm{COP}$, for a RTB refrigerator as a function of the position of the resonant level. c) Efficiency and d) coefficient of performance vs bias voltage for RTB-based devices with $(E_r - E_F)/\Gamma = 1$. In all cases, $k_B T_0/\Gamma = 1$, $(\Gamma_{\mathrm{L}} - \Gamma_{\mathrm{R}})/\Gamma = 0.95$ and the coupling to the gate voltage varies as $C\Gamma/e^2 = 0$ (blue), $5 \cdot 10^{-6}$ (green), and 10^{-5} (red) with a gate voltage given by $eV_g/\Gamma = 3000$. Panels a) and c) are for $T_{\mathrm{R}} - T_{\mathrm{L}} = 0.8\,T_0$ and panels b) and d) are for $T_{\mathrm{R}} - T_{\mathrm{L}} = 0.1\,T_0$. The solid colored lines are the approximate solutions in our weakly nonlinear approach and the dotted colored lines are the fully nonlinear numerical solutions. The dashed black line is the linear response solution and the solid black line is the maximum efficiency/COP predicted from linear response.

before and despite the rather large biases involved, the weakly nonlinear theory gives a rather faithful prediction of the true fully nonlinear behavior.

From the data shown in this section we conclude that nonlinear effects can significantly increase the efficiency of RTB-based thermoelectric engines, first via rectification effects in both the electric and heat currents, and second via capacitive couplings which lift the resonant level.

B. Quantum Point Contact

Nonlinear electric transport through a quantum point contact (QPC) has been discussed theoretically in Ref. [45] and experimentally in Ref. [53]. Linear thermoelectric transport through a QPC has been discussed by Proetto[54], assuming a saddle potential whose transmission spectrum is known exactly[55–57]. Oscillations in the thermopower were in particular reported in Ref. [54], which we reproduce in the inset to the right panel of Fig. 6. Curiously enough, the efficiency of QPC-based thermoelectric devices has not been calculated as of yet, even in the linear regime, and we briefly comment first
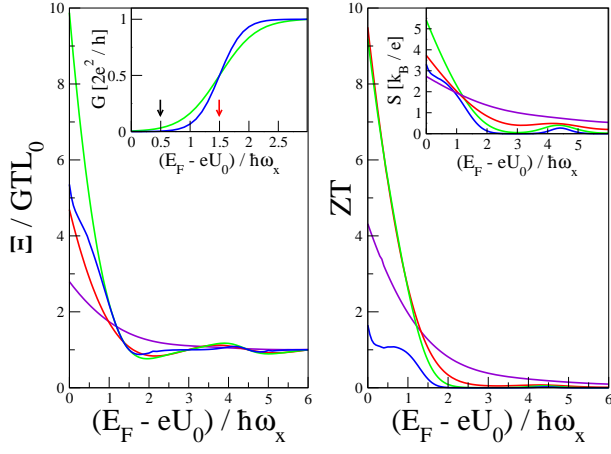
FIG. 6: Left panel: dimensionless ratio Ξ/GTL_0 for a QPC as defined in the text. This ratio is equal to one when the Wiedemann-Franz law is valid. Right panel: dimensionless figure of merit ZT. Left inset: conductance steps for the QPC. The arrows indicate the parameters corresponding to Figs. 8 and 9. Right inset: thermopower S of a QPC. Parameters are $\omega_y/\omega_x = 3$ and $T = 0.10$ (blue), 0.25 (green), 0.50 (red), 1.00 (violet).

on the figure of merit for a QPC. We show in Fig. 6 data for the ratio Ξ/GTL_0 with the Lorenz number $L_0 = \pi^2(k_B/e)^2/3$ as well as $ZT = (GS^2/\kappa)T_0$. When the Wiedemann-Franz law holds, one has $\Xi/GTL_0 = 1$[25]. This is the case when the QPC is well opened and transmits several conductance channels. The Wiedemann-Franz law breaks down when the QPC is closing, and the sharp increase displayed by Ξ/GTL_0 in that regime suggests a reduction in ZT. The right panel of Fig. 6 however indicates an equally sharp increase of ZT reaching almost 10, which is understood once one realizes (see inset) that the increase in Ξ/GTL_0 is accompanied by an increase of S[54], with the net result that ZT increases as the QPC closes. This results in a large efficiency for an almost completely closed QPC operating as a heat engine. This however occurs when the power of the device is exponentially small. Some aspects of nonlinear thermoelectric transport through a QPC have been addressed in Ref. [14]. We here extend these investigations.

The geometry of the QPC is sketched in Fig. 7. Two lateral gates constrain the motion of electrons in a two-dimensional electron gas to a narrow channel. We take these two gates as symmetric and at the same potential and thus consider them together as a single gate. The width of the channel is tuned by the electric potential on the gates. We spatially split the QPC into a left and a right region, each with its own local potential. These two regions are capacitively coupled to one another (capacitance C_0) and to the gate. For physically realistic parameters, an asymmetry in the capacitive coupling of the two regions to the gate has only a small effect on transport, and we therefore consider a symmetric coupling with capacitance C for both the left and right regions of the
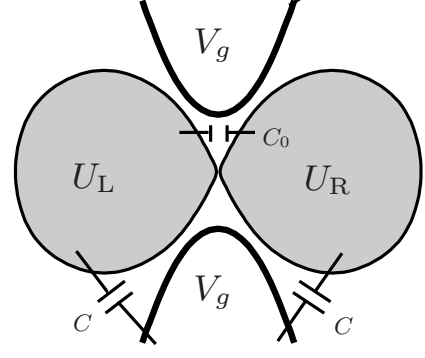

FIG. 7: Sketch of a quantum point contact (QPC) with the local potential in the conductor discretized into a left and a right region. The QPC is electrostatically defined by external gates. The capacitive couplings considered in our theoretical treatment are also indicated.

QPC. We consider a geometrically symmetric QPC with only a few open transmission channels, which we model by a saddle potential[55]

$$eU(x,y) = eU_0 + \frac{1}{2}m\omega_y^2 y^2 - \frac{1}{2}m\omega_x^2 x^2 \,, \qquad (26)$$

with the electron mass m and the electric potential U_0 at the center of the QPC. The transmission probability for each transmission channel is given by[56,57]

$$\mathcal{T}_n(\varepsilon) = \left[1 + e^{-\pi\epsilon_n(\varepsilon)}\right]^{-1} \,, \qquad (27)$$

with

$$\epsilon_n(\varepsilon) = 2\left[\varepsilon + E_F - \hbar\omega_y\left(n + \frac{1}{2}\right) - eU_0\right]\Big/\hbar\omega_x, \quad (28)$$

for $n = 0, 1, 2, \ldots$ The density of states for each channel, $\nu_n(\varepsilon)$, has been calculated in Ref. [58], using a WKB approach. The result exhibits a nonphysical logarithmic singularity close to conductance steps, but retains its validity away from them. We found that electric and heat currents lie on smooth curves except in the immediate vicinity of the steps. The results presented below accordingly use a numerical interpolation to investigate the behavior of QPC-based devices close to conductance steps, whose exact form does not matter, given the very small range where the expression of Ref. [58] breaks down. Other than this interpolation, we use the density of states calculated in Ref. [58].

The QPC has been discretized into two separate regions, and we express the particle injectivity into region $r = \text{L}, \text{R}$ from lead $j = \text{L}, \text{R}$ as[45]

$$\nu_{rj}^p(\varepsilon) = \frac{1}{2}\sum_n \nu_n(\varepsilon)\left[\mathcal{T}_n(\varepsilon) + 2[1 - \mathcal{T}_n(\varepsilon)]\delta_{rj}\right] \,, \quad (29)$$

where the prefactor of 1/2 takes into account that only electrons moving toward the QPC undergo transmission
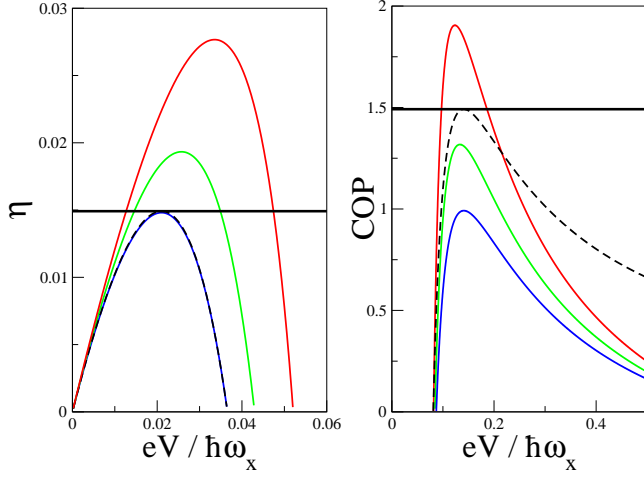
FIG. 8: Left panel: Efficiency of a QPC-based heat engine. Right panel: Coefficient of performance for a QPC-based refrigerator. In both instances the QPC is set at $(E_F - eU_0)/\hbar\omega_x = 1.5$ corresponding to the red arrow in the inset of Fig. 6, $\omega_y/\omega_x = 3$, $eV_g/\hbar\omega_x = 3000$, $k_B T_0/\hbar\omega_x = 0.3$, $T_R - T_L = 0.1\ T_0$, $eV = E_F - \mu_R = \mu_L - \mu_R$, $C_0 = 0$ and $C\hbar\omega_x/e^2 = 0$ (blue), $5 \cdot 10^{-5}$ (green), and 10^{-4} (red).
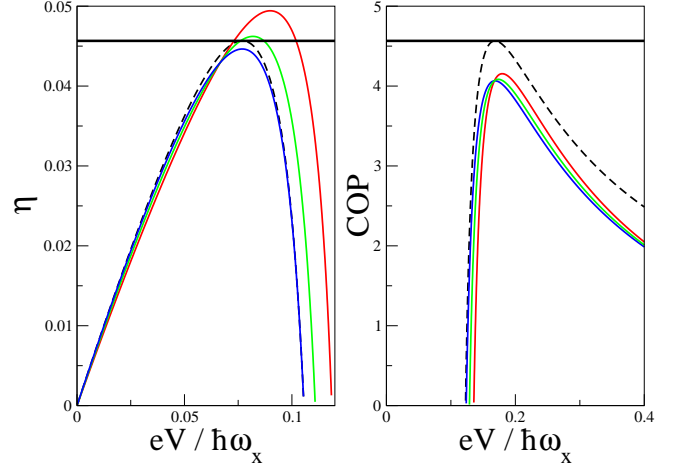
FIG. 9: Left panel: Efficiency of a QPC-based heat engine. Right panel: Coefficient of performance for a QPC-based refrigerator. In both instances the QPC is set at $(E_F - eU_0)/\hbar\omega_x = 0.5$ corresponding to the black arrow in the inset of Fig. 6, $\omega_y/\omega_x = 3$, $eV_g/\hbar\omega_x = 3000$, $k_B T_0/\hbar\omega_x = 0.3$, $T_R - T_L = 0.1\ T_0$, $eV = E_F - \mu_R = \mu_L - \mu_R$, $C_0 = 0$ and $C\hbar\omega_x/e^2 = 0$ (blue), $2 \cdot 10^{-5}$ (green), and $4 \cdot 10^{-5}$ (red).

or reflection. Eq. (10) now becomes

$$\delta q_r \simeq \sum_j D^V_{rj} V_j + \sum_j D^\theta_{rj} \theta_j - DU_r , \qquad (30)$$

where $r = $ L, R labels the left/right regions of the QPC. As in Eq. (23), the injected charge can alternatively be determined by

$$\delta q_r = C(U_r - V_g) + C_0(U_r - U_{\bar{r}}), \qquad (31)$$

where $\bar{r} = L(R)$ when $r = R(L)$. Eqs. (30) and (31) determine the characteristic potentials as

$$u_{rj} = \frac{(C + C_0 + D)D^V_{rj} + C_0 D^V_{\bar{r}j}}{(C + D)(C + 2C_0 + D)}, \quad u_{rg} = \frac{C}{C + D} \quad (32)$$

$$z_{rj} = \frac{(C + C_0 + D)D^\theta_{rj} + C_0 D^\theta_{\bar{r}j}}{(C + D)(C + 2C_0 + D)}, \quad z_{rg} = 0 , \quad (33)$$

which, again, we use to calculate nonlinear transport coefficients in Eqs. (5) and the currents in Eqs. (2). Note that the characteristic potentials now have two labels, one indicating the regions where they are measured, the other one the lead or gate from which they originate.

We present our results for the efficiency of QPC-based devices. For all presented data, we set the chemical potential of the left terminal at E_F and bias the right reservoir to lower electrochemical potential, $\mu_R = E_F - eV$. This corresponds approximately to the situation considered by Whitney[14] (see his Fig.2d), though he chose a Heaviside function for the QPC transmission. Another discrepancy between that work and ours is that we discretize the geometry into two distinct spatial regions,

which accordingly may carry two different local potentials. The data we present all have $C_0 = 0$, as we found that this capacitive coupling does not change the results significantly. Note that for the QPC, only the weakly nonlinear results are shown.

In Fig. 8 we plot both the efficiency of a QPC-based heat engine (left panel) and the coefficient of performance of a QPC-based refrigerator (right panel) as a function of the applied bias voltage when the capacitive coupling C to the gate is varied. Both sets of data correspond to a first transmission channel close to half transmission, $\mathcal{T} \simeq 1/2$, corresponding to the red arrow in the inset of the left panel of Fig. 6. Fig. 8 shows that in the absence of capacitive coupling, the maximal efficiency of the heat engine is slightly lower than one would predict from linear response, in agreement with Ref. [14]. The trend is however reversed once the capacitive coupling C is turned on. Increasing C increases η and even extends the bias range of operation of the heat engine towards higher bias. Turning our attention next to the refrigerator case (right panel in Fig. 8) we note that the discrepancy between linear prediction and nonlinear calculation is stronger, essentially because the refrigerator works at higher bias voltage, therefore larger nonlinear effects may be expected. Increasing C increases the COP for refrigeration just as it increased the power generation efficiency for the QPC when operating as a heat engine. In both instances we see that the nonlinear efficiency significantly exceeds the linear prediction at large enough capacitive coupling C.

Fig. 6 shows a dramatic increase of ZT when the QPC is almost totally closed and its conductance is exponentially small. We investigate whether this region presents

opportunities for thermoelectricity, in particular whether the already large linear efficiency can be further increased by nonlinear effects. In Fig. 9 we plot data similar to those presented in Fig. 8 but obtained at smaller E_F, corresponding to the black arrow in the inset of the left panel of Fig. 6. We see that nonlinearities influence thermodynamic performances only weakly in that regime, essentially because currents are exponentially small.

VII. CONCLUSIONS

We have presented a weakly nonlinear theory of coupled heat and electric transport in micro- and nanoscopic devices. On those scales it is of central importance that local electrostatic potentials are correctly taken care of, in order to preserve fundamental conservation laws such as gauge invariance and current conservation. We described how this can be done and listed sum rules ensuring that these conservation laws are preserved. Most works on thermoelectricity use the linear figure of merit ZT as a metric for thermodynamic performance – the larger ZT, the better the efficiency of a heat engine and the coefficient of performance of a refrigerator. We

pointed out that the predictive power of ZT is no longer guaranteed when nonlinear effects kick in, and directly investigated the thermodynamic efficiency and coefficient of performance, including nonlinear contributions to the heat and electric currents. We investigated devices based on two microconstrictions, a quantum point contact and a resonant tunneling barrier. In both instances we found situations where the true efficiency significantly differs from that predicted by linear response, as well as cases where thermodynamic efficiencies strongly depend on the strength of these nonlinearities – not only on the magnitude of the applied temperature and voltage biases, but also on the capacitive coupling to external electrostatically charge regions. Most importantly, we found that the weakly nonlinear theory faithfully captures the true nonlinear behavior of micro- and nanoscopic devices. We believe this opens the way to more general investigations of thermoelectricity in quantum microdevices which until now have often been limited to numerical investigations.

[1] S. Chu and A. Majumdar, Nature **488**, 294 (2012).

[2] G. Mahan, B. Sales, and J. Sharp, Phys. Today **50**, vol. 3, 42 (1997).

[3] A. Majumdar, Science **303**, 777 (2004).

[4] G.J. Snyder and E.S. Toberer, Nat. Materials **7**, 105 (2008).

[5] A. Shakouri, Ann. Rev. Mat. Res. **41**, 399 (2011).

[6] M.S. Dresselhaus, G. Chen, M.Y. Tang, R. Yang, H. Lee, D. Wang, Z. Ren, J.-P. Fleurial, and P. Gogna, Adv. Mater. **19**, 1043 (2007).

[7] G.D. Mahan and J.O. Sofo, PNAS **93**, 7436 (1996).

[8] F. Giazotto, T.T. Heikkilä, A. Luukanen, A.M. Savin, and J.P. Pekola, Rev. Mod. Phys. **78**, 217 (2006).

[9] R. Scheibner, M. König, D. Reuter, A.D.Wieck, C. Gould, H. Buhmann, and L.W. Molenkamp, New J. Phys. **10**, 083016 (2008).

[10] B. Sothmann, R. Sanchez, A.N. Jordan, and M. Büttiker, Phys. Rev. B **85**, 205301 (2012).

[11] J. Matthews, D. Sanchez, M. Larsson, and H. Linke, Phys. Rev. B **85**, 205309 (2012).

[12] H.L. Edwards, Q. Niu, G.A. Georgakis, and A.L. de Lozanne, Phys. Rev. B **52**, 5714 (1995).

[13] J.R. Prance, C.G. Smith, J.P. Griffiths, S.J. Chorley, D. Anderson, G.A.C. Jones, I. Farrer, and D.A. Ritchie, Phys. Rev. Lett. **102**, 146602 (2009).

[14] R.S. Whitney, arXiv:1208.6130.

[15] T.E. Humphrey, R. Newbury, R.P. Taylor, and H. Linke, Phys. Rev. Lett. **89**, 116801 (2002).

[16] R. Sanchez and M. Büttiker, Phys. Rev. B **83**, 085428 (2011).

[17] B. Muralidharan and M. Grifoni, Phys. Rev. B **85**, 155423 (2012).

[18] P. Reddy, S.Y. Jang, R.A. Segalman, and A. Majumdar, Science **315**, 1568 (2007).

[19] K. Baheti, J.A. Malen, P. Doak, P. Reddy, S.Y. Jang, T.D. Tilley, A. Majumdar, and R.A. Segalman, Nano Lett. **8**, 715 (2008).

[20] J.P. Bergfield and C.A. Stafford, Nano Lett. **9**, 3072 (2009).

[21] D. Nozaki, H. Sevincli, W. Li, R. Gutiérrez, and G. Cuniberti, Phys. Rev. B **81**, 235406 (2010).

[22] O. Entin-Wohlman, Y. Imry, and A. Aharony, Phys. Rev. B **82**, 115314 (2010).

[23] S.K. Yee, J.A. Malen, A. Majumdar, and R.A. Segalman, Nano Lett. **11**, 4089 (2011).

[24] B.K. Nikolić, K.K. Saha, T. Markussen, and K.S. Thygesen, J. Comput. Electron. **11**, 78 (2012); K.K. Saha, T. Markussen, K.S. Thygesen, and B.K. Nikolić, Phys. Rev. B **84**, 041412(R) (2011).

[25] N.W. Ashcroft and N.D. Mermin, *Solid State Physics*, Saunders (Philadelphia, 1976).

[26] R.W. Hill, C. Proust, L. Taillefer, P. Fournier, and R.L. Greene, Nature **414**, 711 (2001).

[27] P. Coleman, J.B. Marston, and A.J. Schofield, Phys. Rev B **72**, 245111 (2005).

[28] D. Podolsky, A. Vishwanath, J. Moore, and S. Sachdev, Phys. Rev B **75**, 014520 (2007).

[29] M.G. Vavilov and A.D. Stone, Phys. Rev. B **72**, 205107 (2005).

[30] V. Balachandran, R. Bosisio, and G. Benenti, Phys. Rev. B **86**, 035433 (2012).

[31] Z. Jiang and V. Chandrasekhar, Phys. Rev. B **72**, 020502(R) (2005).

[32] P. Virtanen and T.T. Heikkilä, Appl. Phys. A **89**, 625 (2007).

[33] Ph. Jacquod and R.S. Whitney, Europhys. Lett. **91**, 67009

(2010).

[34] A.S. Dzurak, C.G. Smith, L. Martin-Moreno, M. Pepper, D.A. Ritchie, G.A.C. Jones, and D.G. Hasko, J. Phys.: Condens. Matter **5**, 8055 (1993).

[35] I. Shorubalko, H. Q. Xu, I. Maximov, P. Omling, L. Samuelson, and W. Seifert, Appl. Phys. Lett. **79**, 1384 (2001).

[36] M. Terraneo, M. Peyrard, and G. Casati, Phys. Rev. Lett. **88**, 094302 (2002).

[37] D. Sanchez and M. Büttiker, Phys. Rev. Lett. **93**, 106802 (2004).

[38] B. Spivak and A. Zyuzin, Phys. Rev. Lett. **93**, 226801 (2004).

[39] A.V. Andreev and L.I. Glazman, Phys. Rev. Lett. **97**, 266806 (2006).

[40] J. Meair and Ph. Jacquod, J. Phys.: Condens. Matter **24**, 272201 (2012).

[41] A.N. Grigorenko, P.I. Nikitin, D.A. Jelski, and T.F. George, Phys. Rev. B **42**, 7405 (1990).

[42] I.O. Kulik, J. Phys. condens. Matter **6**, 9737 (1994).

[43] J.K. Freericks and V. Zlatic, Condens. Matter Phys. **9**, 603 (2006).

[44] M. Zebarjadi, K. Esfarjani, and A. Shakouri, Appl. Phys. Lett. **91**, 122104 (2007).

[45] T. Christen and M. Büttiker, Europhys. Lett. **35**, 523 (1996); M. Büttiker and T. Christen, in *Mesoscopic Electron Transport*, NATO ASI Series E, Eds. L. Kouwenhoven, G. Schoen, and L. Sohn, Kluwer (1997).

[46] E.N. Bogachek, A.G. Scherbakov, and U. Landman, Phys. Rev. B **60**, 11678 (1999).

[47] M.A. Çipiloğlu, S. Turgut, and M. Tomak, Phys. Status Solidi B **241**, 2575 (2004).

[48] D. Sanchez and R. Lopez, arXiv:1209.1264.

[49] Eq. (18) has to be modified when time-reversal symmetry is broken by a field ϕ in systems where the thermopower is not symmetric in ϕ, $S(\phi) \neq S(-\phi)$. Such systems include Andreev systems, see Ref. [33], and multiterminal devices, see: G. Benenti, K. Saito, and G. Casati, Phys. Rev. Lett. **106**, 230602 (2011).

[50] P.N. Butcher, J. Phys.: Condens. Matter **2**, 4869 (1990).

[51] Ph. Jacquod, R.S. Whitney, J. Meair, and M. Büttiker, Phys. Rev. B **86**, 155118 (2012).

[52] In this work we define the thermopower as $S \equiv (\Delta V/\Delta T)_{I=0}$ whereas in other works it is often defined with an additional minus sign, $S \equiv -(\Delta V/\Delta T)_{I=0}$.

[53] N.K. Patel, J.T. Nicholls, L. Martín-Moreno, M. Pepper, J.E.F. Frost, D.A. Ritchie, and G.A.C. Jones, Phys. Rev. B **44**, 13549 (1991).

[54] C.R. Proetto, Phys. Rev. B **44**, 9096 (1991).

[55] M. Büttiker, Phys. Rev. B **41**, 7906 (1990).

[56] J.N.L. Connor, Mol. Phys. **15**, 37 (1968).

[57] W.H. Miller, J. Chem. Phys. **48**, 1651 (1968).

[58] M. H. Pedersen, S. A. van Langen, and M. Büttiker, Phys. Rev. B **57**, 1838 (1997)

[59] R.S. Whitney, arXiv:1211.4737.